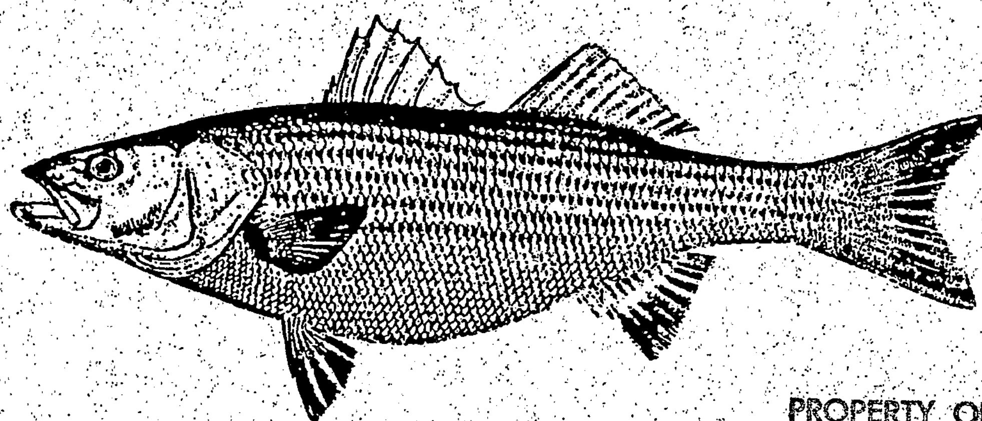
WRINT-DFG-Exhibit 3

# A MODEL FOR EVALUATING THE IMPACTS OF FRESHWATER OUTFLOW AND EXPORT ON STRIPED BASS IN THE SACRAMENTO-SAN JOAQUIN ESTUARY





David W. Kohlhorst
Donald E. Stevens
Lee W. Miller

Bay-Delta and Special Water Projects Division
California Department of Fish and Game
4001 North Wilson Way
Stockton, California 95205-2424

Entered by the California Department of Fish and Game
for the State Water Resources Control Board 1992
Water Rights Phase of the Bay-Delta Estuary Proceedings

# TABLE OF CONTENTS

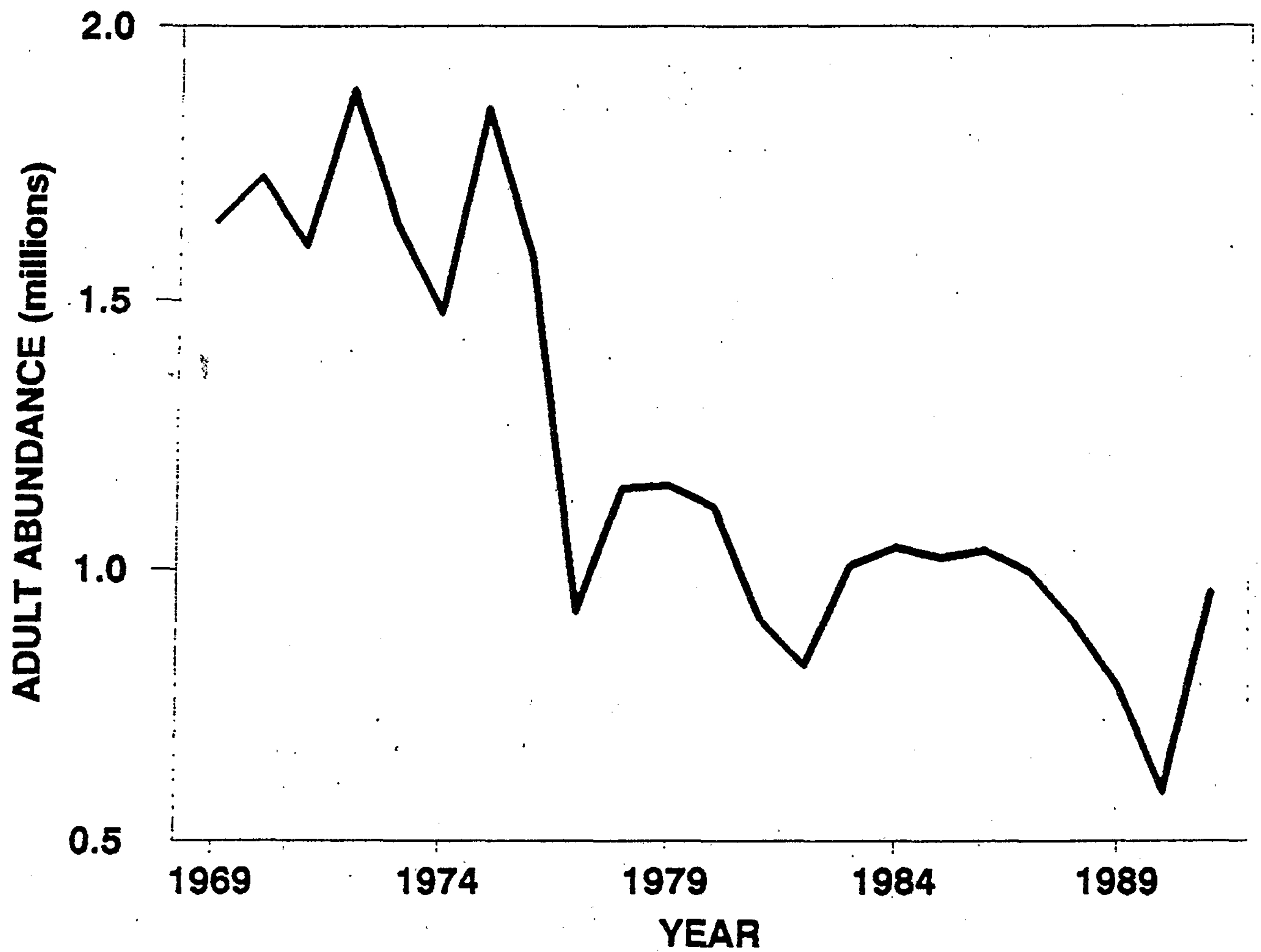



Figure 1. Trend in legal-sized striped bass abundance in the Sacramento-San Joaquin Estuary, 1969-1991.

months from late summer through winter of 21-150 mm fish at the State Water Project (SWP) and Central Valley Project (CVP) export pumps in the south Delta (Table 1) (DFG 1992). These post-yoy losses have been estimated to range from less than 200,000 bass in 1983 to almost 22 million fish in 1974 (Figure 3). The loss estimates assume size-dependent predation losses in the SWP's Clifton Court Forebay beginning in 1971 which range from 93% for 21-25-mm bass to 3% for 141-150-mm fish (Table 2). Size-dependent predation losses at the Federal CVP fish screening facility where there is no forebay (and at the SWP facility before 1971 when a large predator population had developed) were scaled, for the same size range, from 17% to 1% (Table 2). For consistency, the Clifton Court Forebay predation curve is that used in the Four Pumps Mitigation Agreement. However, this curve appears to underestimate predation mortality when compared to results of experiments conducted with yoy striped bass (mean fork length from 47 to 56 mm) which found loss rates in the forebay of 94% in July, 1984 and 70% in August, 1986 (Kano 1985, 1986).

The magnitude of post-yoy index losses at the water export pumps is potentially affected by three readily identifiable factors: (1) the abundance of young bass; (2) the magnitude of water exports; and (3) Delta outflow, because it influences distribution of the young fish and their vulnerability to entrainment with exported water. For the purpose of evaluating the influence of water exports and outflow, the effect of young bass abundance can be removed by dividing post-yoy losses by the yoy index to produce a loss rate index which, conceptually, is similar to "fraction of the population removed" and is expressed as export loss per yoy index unit. This loss rate index has increased dramatically in recent years, from low values in the tens of thousands in the 1960s when only the CVP was exporting water from the Delta to over one million in 1987 and 1989 when both projects exported large amounts of water (Figure 4).

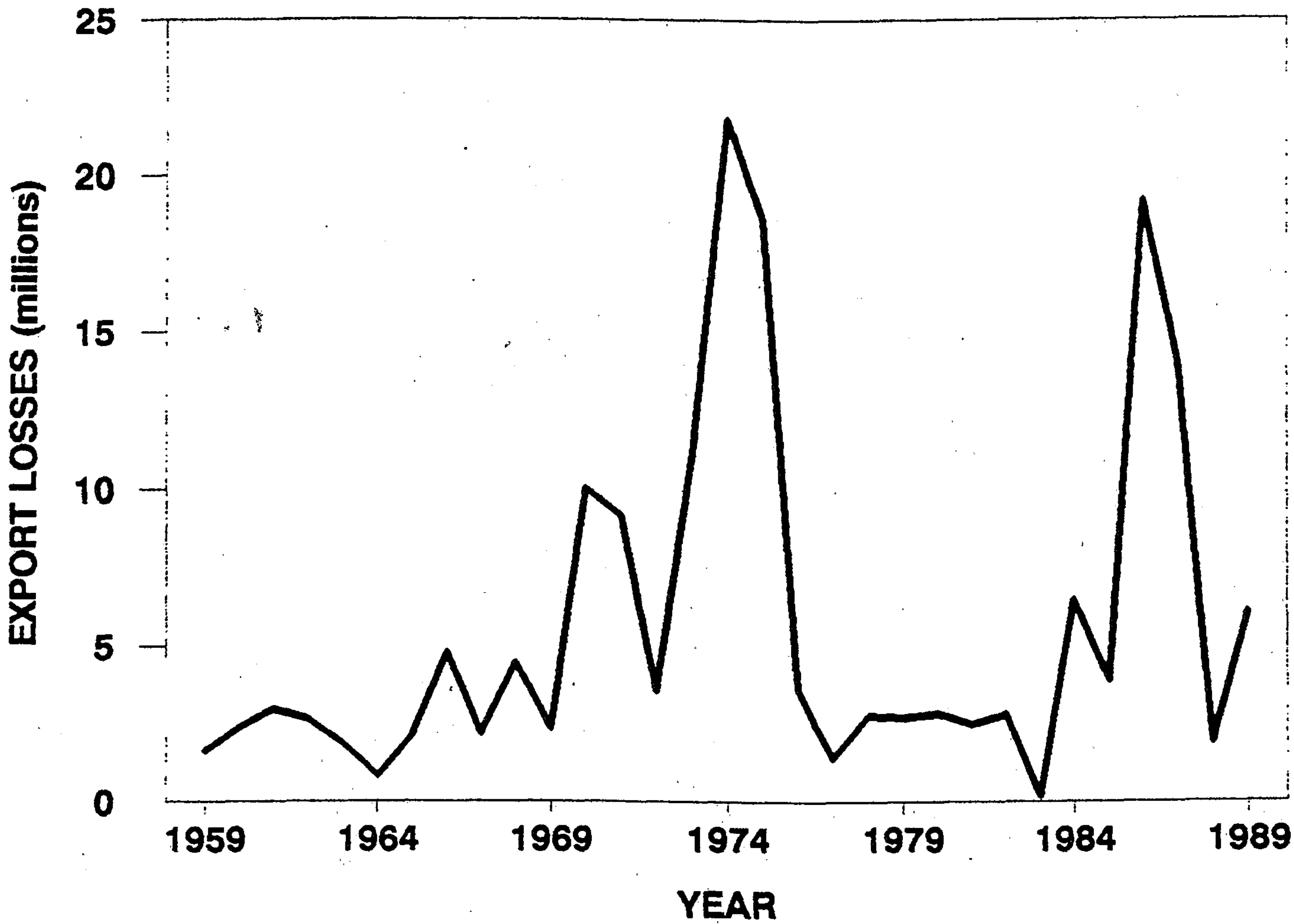



Figure 3. Trend in estimated losses to Central Valley Project and State Water Project export pumping of 21-150 mm striped bass after the time when the young-of-the-year index is set. Estimates assume size-dependent predation mortality in Clifton Court Forebay and at the CVP fish screens.

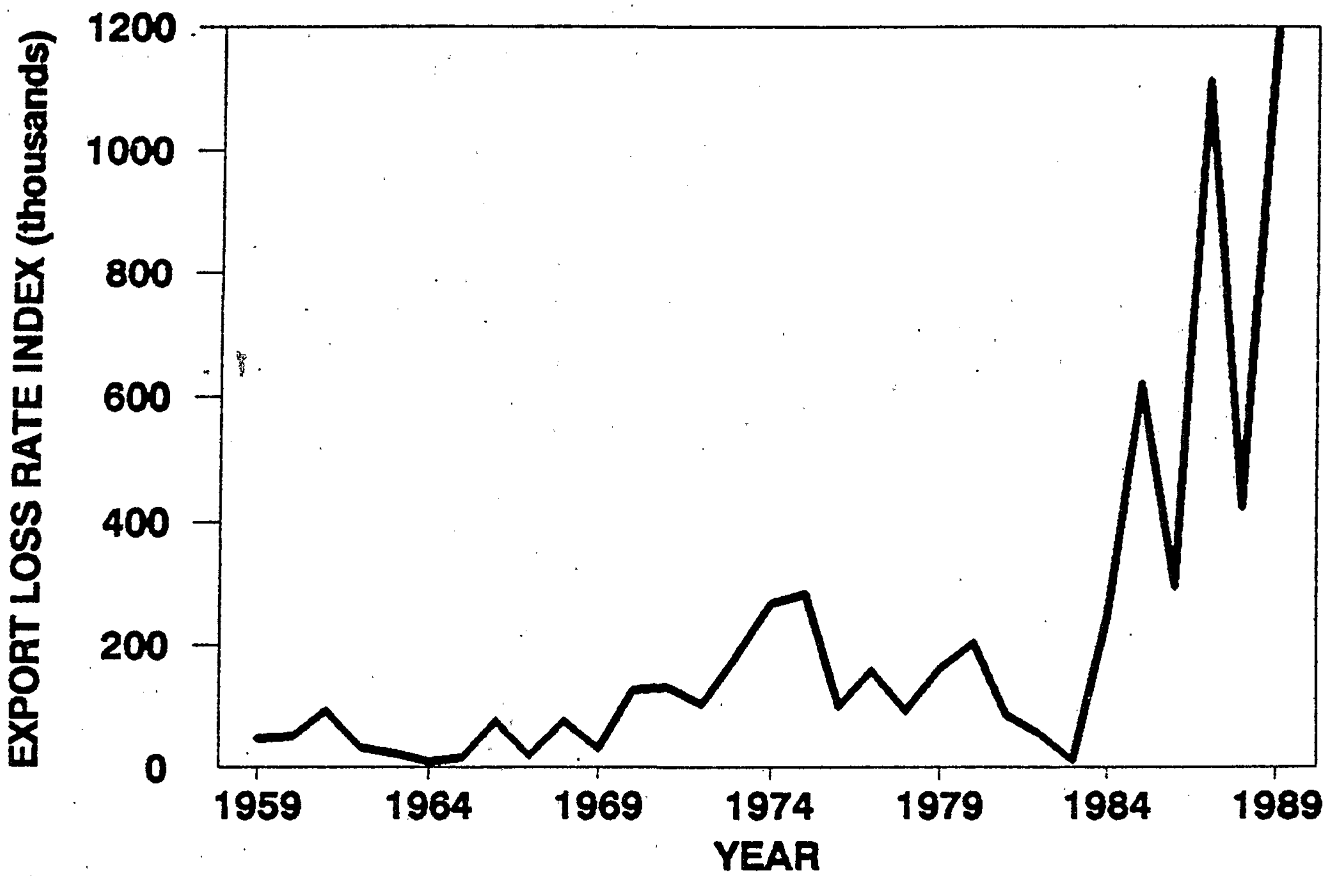



Figure 4. Trend in estimated loss rate of 21-150 mm striped bass to Central Valley Project and State Water Project export pumping after the time when the young-of-the-year index is set. Loss rate is the estimated export loss divided by the young-of-the-year index and represents the number of young bass lost per index unit.

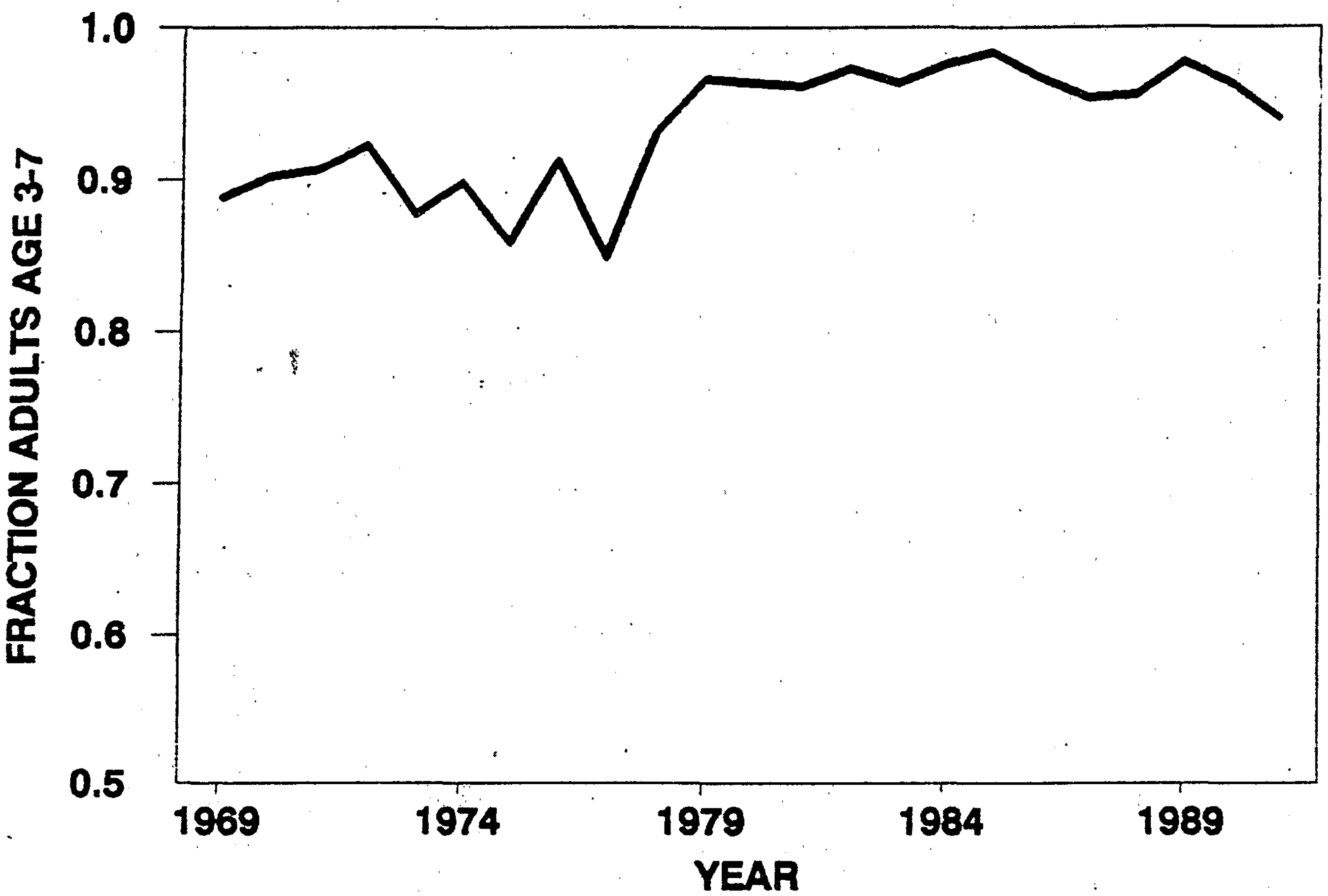



Figure 5. Proportion of the legal-sized striped bass abundance estimate that is age 3-7.

Table 4. Results of correlation analysis between wild adult striped bass abundance (without hatchery-produced fish) and weighted mean yoy abundance index, weighted mean post-yoy losses, and weighted mean post-yoy loss rate 3-7 years earlier.

| | ADULTS | $LOG_{10}$(ADULTS) |
|---|---|---|
| MEAN YOY | 0.775 | 0.756 |
| $LOG_{10}$(MEAN YOY) | 0.742 | 0.723 |
| MEAN LOSSES | -0.263 | -0.282 |
| $LOG_{10}$(MEAN LOSSES) | -0.186 | -0.210 |
| MEAN LOSS RATE | -0.619 | -0.679 |
| $LOG_{10}$(MEAN LOSS RATE) | -0.727 | -0.747 |

results from initially strong year classes that experience only small late summer through winter losses to export pumping. We decided to use the yoy abundance index in combination with loss rate rather than losses in the final equation to describe the effects of these variables on adult striped bass abundance. The model with loss rate is more straightforward because it allows evaluation of post-yoy index water management scenarios that are not dependent on the yoy index. The equation

LEGAL-SIZED ADULTS = 18940 WEIGHTED MEAN YOY INDEX - 446608 LOG(WEIGHTED MEAN LOSS RATE) + 2960840

explains 71% of the variability in adult striped bass abundance (Figure 6).

## VERIFICATION OF THE PREDICTABILITY OF ADULT STRIPED BASS ABUNDANCE FROM YOUNG STRIPED BASS ABUNDANCE AND SUBSEQUENT ENTRAINMENT LOSSES

Other data and methods were explored for the purpose of evaluating the reasonableness of the results relating adult striped bass abundance to young bass abundance and entrainment losses.

### Discriminant Analysis

Stepwise discriminant analysis with the same linear and log-transformed variables employed in the above regression analysis was used to assign the annual adult population estimate to one of two groups, high abundance (>1.4 million) or low abundance (<1.2 million). A jackknife validation procedure (Dixon 1988, p 337; Johnson and Wichern 1988, p 498) classified each year into a group based on classification functions computed from all years except the year being classified. Jackknife discriminant analysis was 100% successful at assigning each year's adult population estimate to the proper group with classification functions which selected weighted mean yoy, weighted mean export loss, and log(weighted mean export loss) as significant variables (Table 6). Five replications of an analysis which randomly split the data set and used the classification functions developed from one subset to classify the years in the other subset resulted in a high proportion of correct classifications in the test subsets (Table 6).

Thus, this approach provides strong support for our model.

Analysis with Ages 3, 4, and 5

Petersen population estimates are available for individual age groups up to age 7 (Table 3) so that the relationship of each age group to its abundance in the first summer of life and subsequent first-year entrainment losses can be explored. We chose to examine this relationship for recruits (ages 3 and 4) and age 5, which is the age at which most females become sexually mature and, thus, fully vulnerable to capture by our tagging program during the spring spawning migration.

Stepwise regression of estimated abundance at each age and consecutive combinations of ages on yoy index, export losses, and loss rate with appropriate lags (weighted means over the appropriate years for combinations of ages) yielded results that were generally consistent with the analysis using total adult abundance (Table 7). In all cases (except for age 4), yoy index and export losses produced the "best" model (highest $R^2$ and including all independent variables allowed to enter by the stepwise process), explaining from 42% to 65% of the variance in abundance of individual or combinations of ages. Loss rate was also related to abundance, but explained much of the same variance as the yoy index and was removed from the model when yoy entered.

The results with the individual ages generally support our model.

Table 7. Results of stepwise regression of wild age 3-5 striped bass abundance (without hatchery-produced fish) on the yoy abundance index (YOY), post-yoy losses (LOSSES), and post-yoy loss rate (LOSS RATE). Combinations of ages are regressed on weighted means of the independent variables with appropriate time lags. Weighting factors are age-class abundance relative to age 3 (Table 3). Values are coefficients of determination ($R^2$) expressed as percentages. The $R^2$ value for the final model selected by stepwise regression is underlined.

| Independent Variables | Age 3 | Age 4 | Age 5 | Age 3 & 4 | Age 4 & 5 | Age 3-5 |
|---|---|---|---|---|---|---|
| YOY | 27 | 6 | 27 | 38 | 21 | 52 |
| LOSSES | 2 | 20 | 5 | 4 | 12 | 4 |
| LOSS RATE | 17 | 18 | 21 | 28 | <u>28</u> | 34 |
| YOY & LOSSES | <u>42</u> | <u>33</u> | <u>44</u> | <u>54</u> | 42 | <u>65</u> |
| YOY & LOSS RATE | 33 | 19 | 36 | 47 | 37 | 61 |

Table 8. Results of stepwise regression of wild adult striped bass abundance (without hatchery-produced fish) on the weighted mean yoy abundance index (WTMNYOY), weighted mean post-yoy yearling equivalent losses (WTMNYELOSS), and mean weighted post-yoy yearling equivalent loss rate (WTMNYELOSSRATE) 3-7 years earlier. Weighting factors are age-class abundance relative to age 3 (Table 3). Results with linear and log-transformed values of adult abundance are presented. Values in the table are coefficients of determination ($R^2$) expressed as percentages. The $R^2$ value for the final model selected by stepwise regression is underlined.

| Independent Variables | ADULTS | $LOG_{10}$(ADULTS) |
|---|---|---|
| WTMNYOY | 60 | <u>57</u> |
| WTMNYELOSS | 18 | 15 |
| WTMNYELOSSRATE | 43 | 42 |
| WTMNYOY & WTMNYELOSS | <u>67</u> | 63 |
| WTMNYOY & WTMNYELOSSRATE | 61 | 58 |

Table 9. Catch-per-effort index of striped bass abundance developed from catches of legal-sized fish during annual spring tagging in the western Delta and in the Sacramento River near Clarksburg. Annual effort is four boat-months of gill netting and 36 trap-months of trapping. Traps were not fished in 1977 and 1978 and were fished at other locations in 1981 and after 1989.

| Year | Catch-per-Effort Index |
|---|---|
| 1969 | 25447 |
| 1970 | 19623 |
| 1971 | 23207 |
| 1972 | 19812 |
| 1973 | 19898 |
| 1974 | 15075 |
| 1975 | 10691 |
| 1976 | 11930 |
| 1977 | Missing |
| 1978 | Missing |
| 1979 | 13249 |
| 1980 | 7394 |
| 1981 | Missing |
| 1982 | 6077 |
| 1983 | 6532 |
| 1984 | 5919 |
| 1985 | 8805 |
| 1986 | 9257 |
| 1987 | 9436 |
| 1988 | 9107 |
| 1989 | 11906 |



C-043017

Table 11. Results of detrending adult abundance, weighted mean yoy index, weighted mean export losses, and weighted mean loss rate by differencing so that $x_i = x_i - x_{i-1}$, where i = year.

*Time trend: variable regressed on year*

| Variable | Original Data Slope | Original Data $r^2$ | Detrended Data Slope | Detrended Data $r^2$ |
|---|---|---|---|---|
| ADULTS | -47513 | 0.74 | 7335 | 0.02 |
| WTMNYOY | -1.383 | 0.80 | 0.018 | 0.00 |
| WTMNLOSS | 27357 | 0.01 | -8175 | 0.00 |
| WTMNLOSSRATE | 7471 | 0.48 | 1513 | 0.05 |
| ***Relationship with Adults*** | | | | |
| WTMNYOY | 27684 | 0.61 | -18145 | 0.05 |
| WTMNLOSS | -0.0533 | 0.07 | -0.0710 | 0.08 |
| WTMNLOSSRATE | -3.157 | 0.38 | -1.283 | 0.02 |

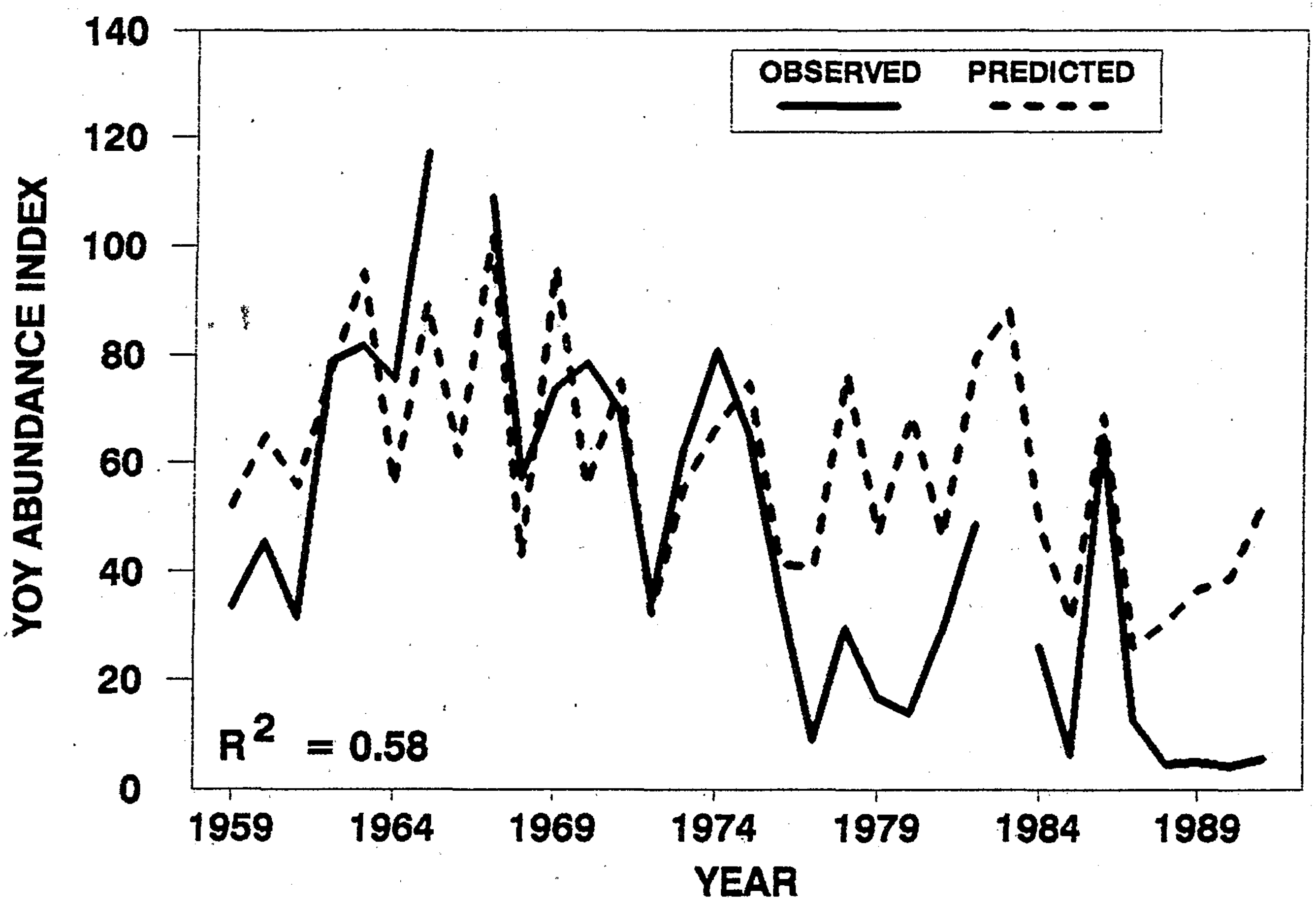



Figure 7. Observed and predicted striped bass young-of-the-year indices from 1959 to 1991. The following prediction equations are based on 1959-1976 data only:
DELTA INDEX = 292.332 LOG(APRIL-JULY OUTFLOW) - 34.866 (LOG(APRIL-JULY OUTFLOW))$^2$ - 0.00561 APRIL-JULY DIVERSIONS - 534.5475
SUISUN INDEX = 46.680 LOG(APRIL-JULY OUTFLOW) - 159.077.
For the April-July period, diversions = exports + 3108.

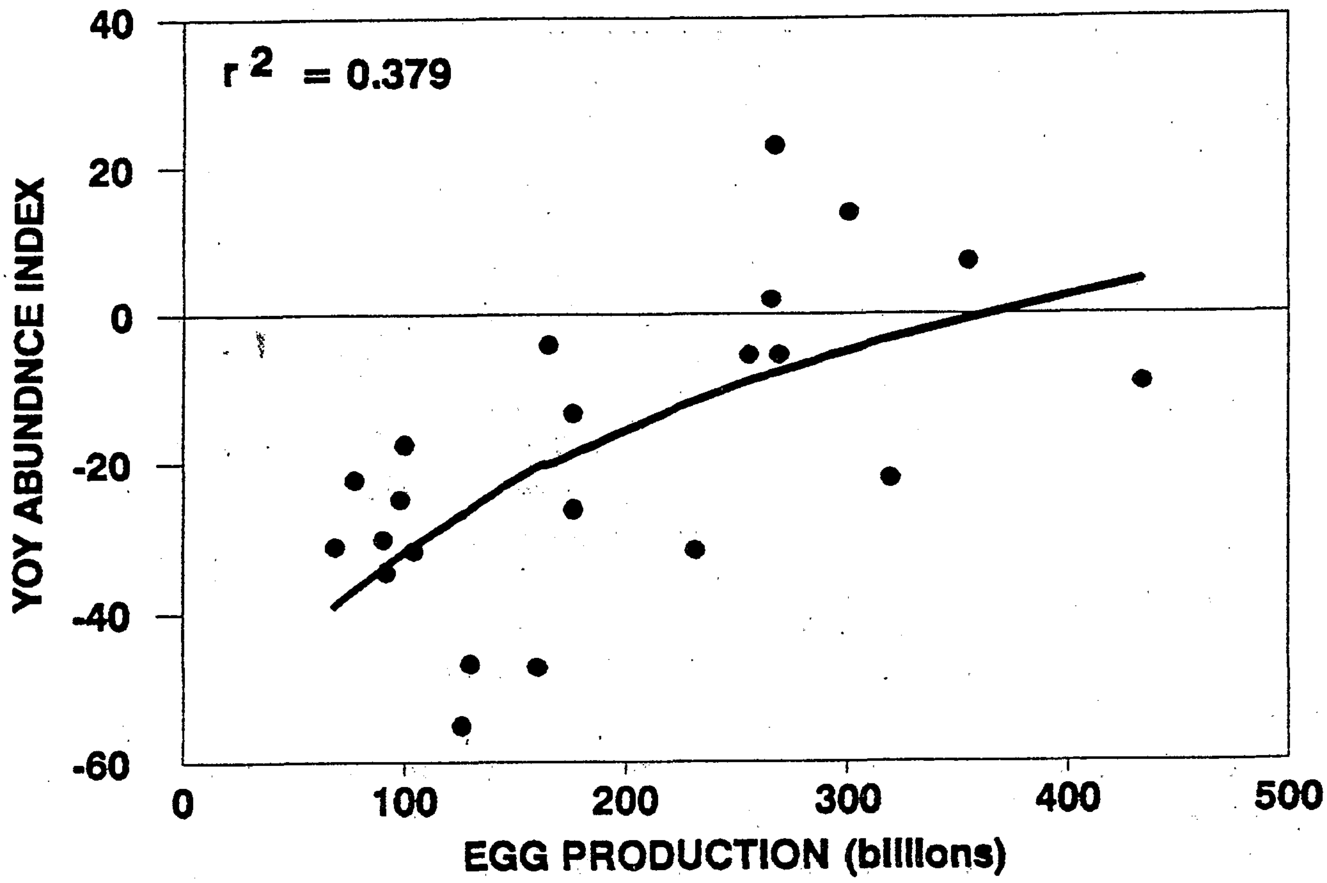



Figure 8. Stock-recruit relationship for striped bass in the Sacramento-San Joaquin Estuary based on the residual young-of-the-year index (after removing the effect of flows and diversions) and estimated egg production (<u>in billions</u>) from the Petersen population estimate and age-specific fecundity estimates. The predictive equation is:

RESIDUAL YOUNG-OF-THE-YEAR = 1/(0.0095 + (2.59/EGGS)) - 60.

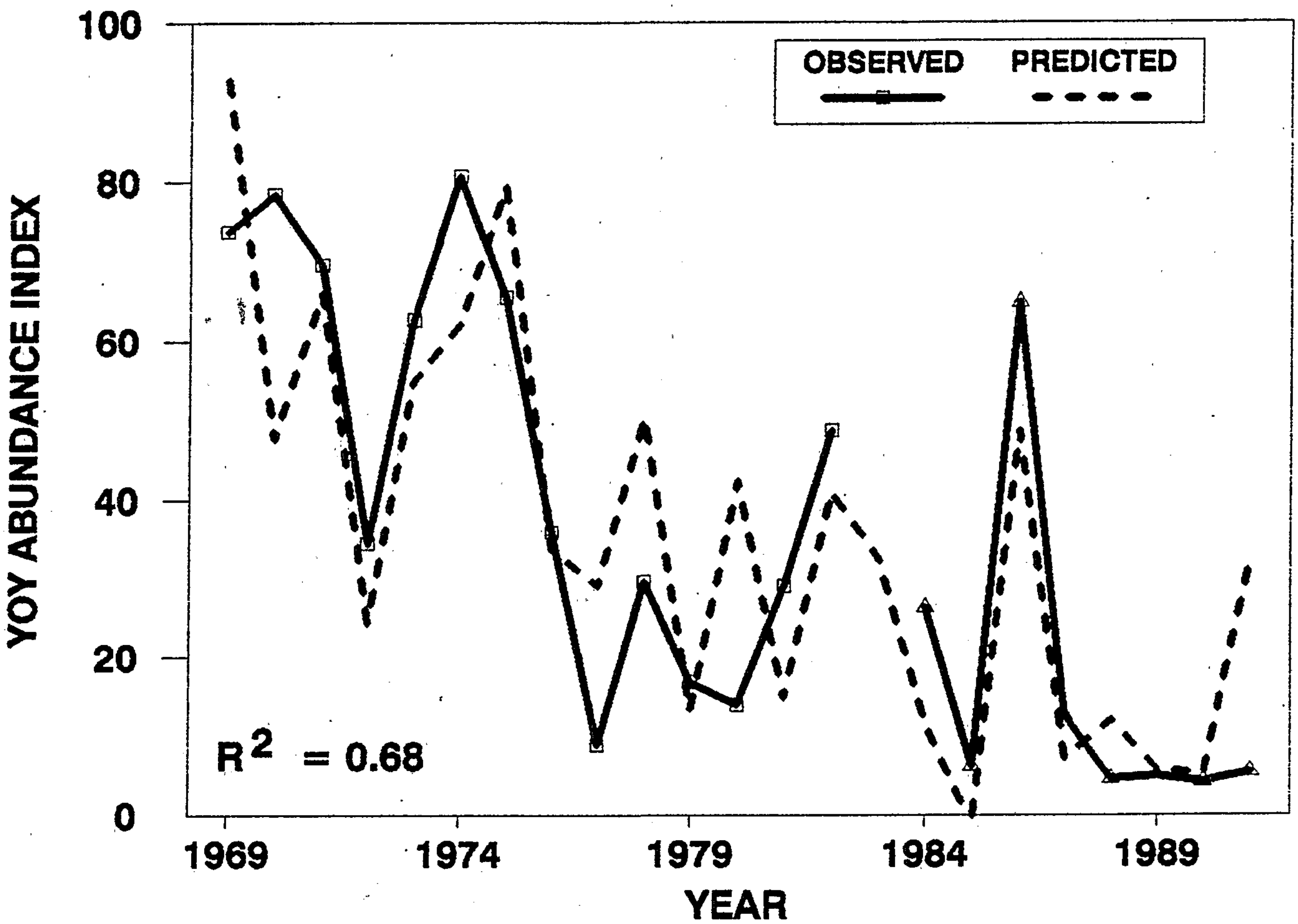



Figure 9. Observed and predicted young-of-the-year indices where predicted values are based on April-July outflow and diversions (Figure 7) and the stock-recruit relationship (Figure 8).

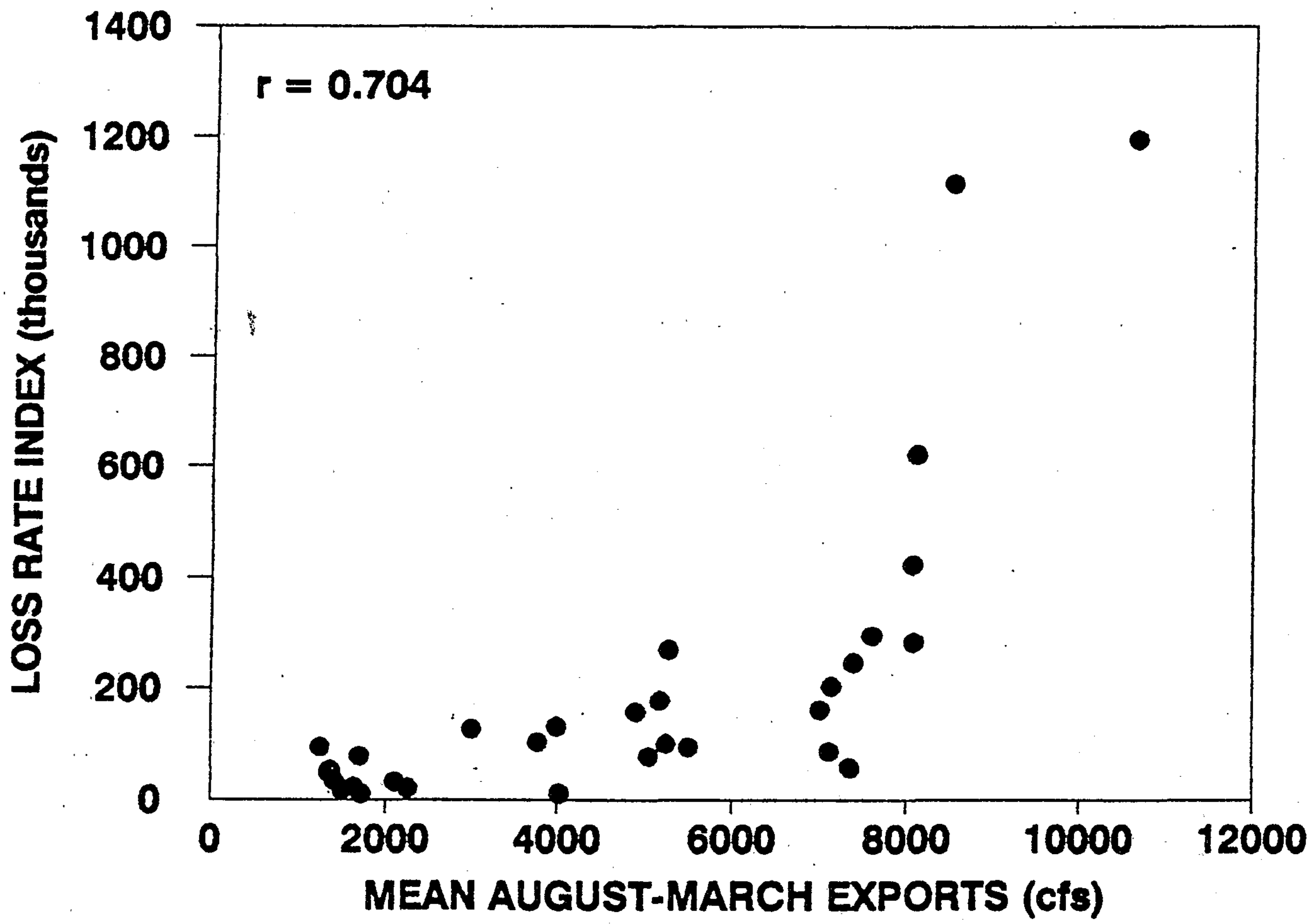



Figure 10. Scatterplot of export loss rate and mean August-March exports from 1959-1989.

Table 13. Correlation coefficients of the residuals from the regression of log(loss rate) on August-March exports with all monthly combinations of August to March outflows.

| MONTH | CORRELATION COEFFICIENT |
|---|---|
| Aug | -0.484 |
| Sep | -0.491 |
| Oct | -0.399 |
| Nov | -0.499 |
| Dec | -0.570 |
| Jan | -0.408 |
| Feb | -0.275 |
| Mar | -0.228 |
| Aug-Sep | -0.495 |
| Sep-Oct | -0.478 |
| Oct-Nov | -0.520 |
| Nov-Dec | -0.571 |
| Dec-Jan | -0.532 |
| Jan-Feb | -0.383 |
| Feb-Mar | -0.283 |
| Aug-Oct | -0.492 |
| Sep-Nov | -0.539 |
| Oct-Dec | -0.583 |
| Nov-Jan | -0.550 |
| Dec-Feb | -0.478 |
| Jan-Mar | -0.366 |
| Aug-Nov | -0.542 |
| Sep-Dec | -0.593 |
| Oct-Jan | -0.567 |
| Nov-Feb | -0.504 |
| Dec-Mar | -0.447 |
| Aug-Dec | -0.596 |
| Sep-Jan | -0.580 |
| Oct-Feb | -0.520 |
| Nov-Mar | -0.471 |
| Aug-Jan | -0.586 |
| Sep-Feb | -0.536 |
| Oct-Mar | -0.486 |
| Aug-Feb | -0.546 |
| Sep-Mar | -0.500 |
| Aug-Mar | -0.508 |

Table 14. Results of stepwise regression of log(loss rate) on mean August-December outflow (A-D OUT)and mean August-March exports (A-M EXP). Values are coefficients of determination ($R^2$) expressed as percentages. The $R^2$ value for the final model selected by stepwise regression is underlined.

| Independent Variables | Log(Loss Rate) |
|---|---|
| A-D OUT | 29 |
| A-M EXP | 63 |
| A-D OUT & A-M EXP | <u>77</u> |



C-043024

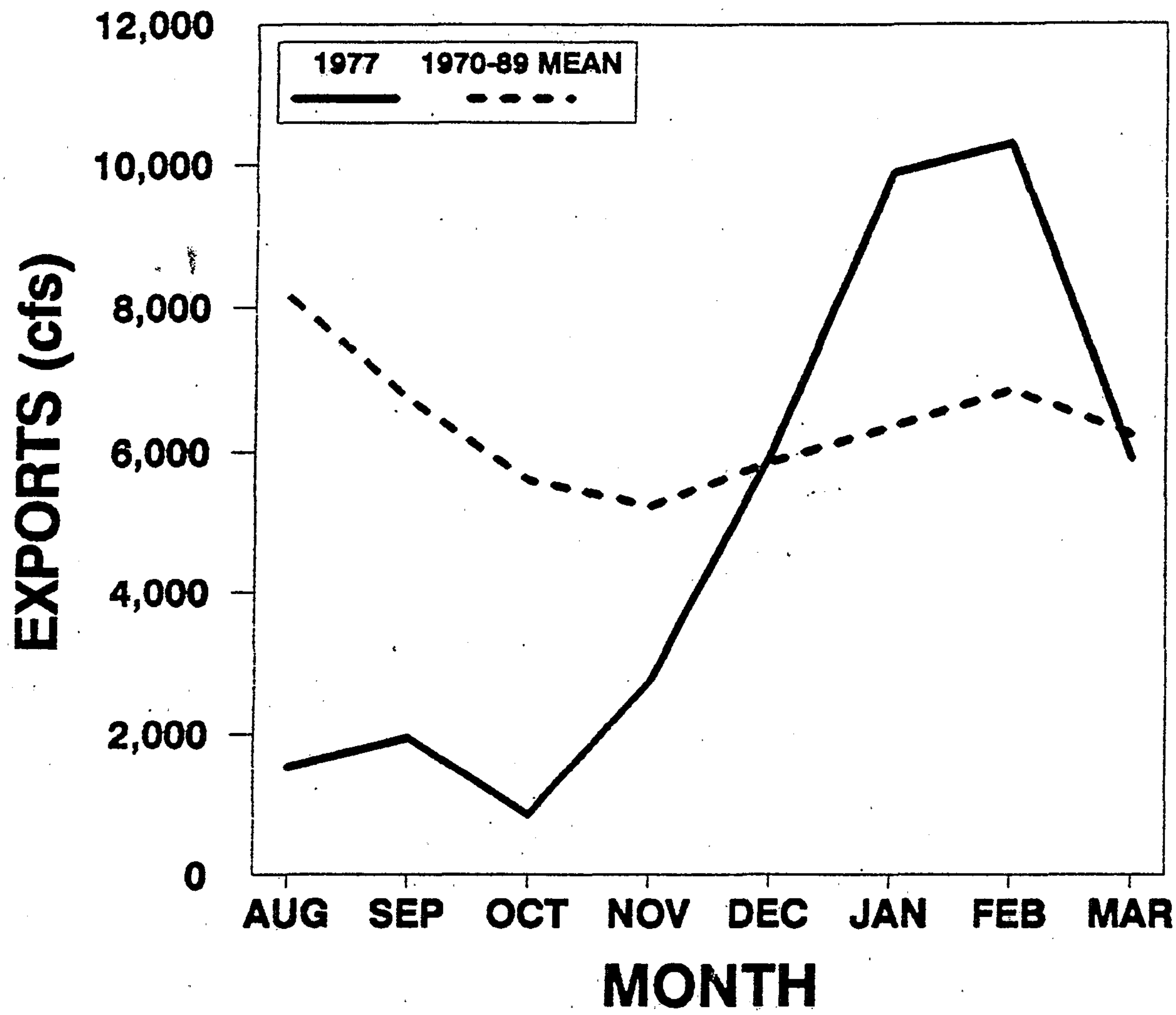



Figure 13. Comparison of mean monthly water exports by the CVP and SWP in 1977 with mean monthly exports in 1970-1989.

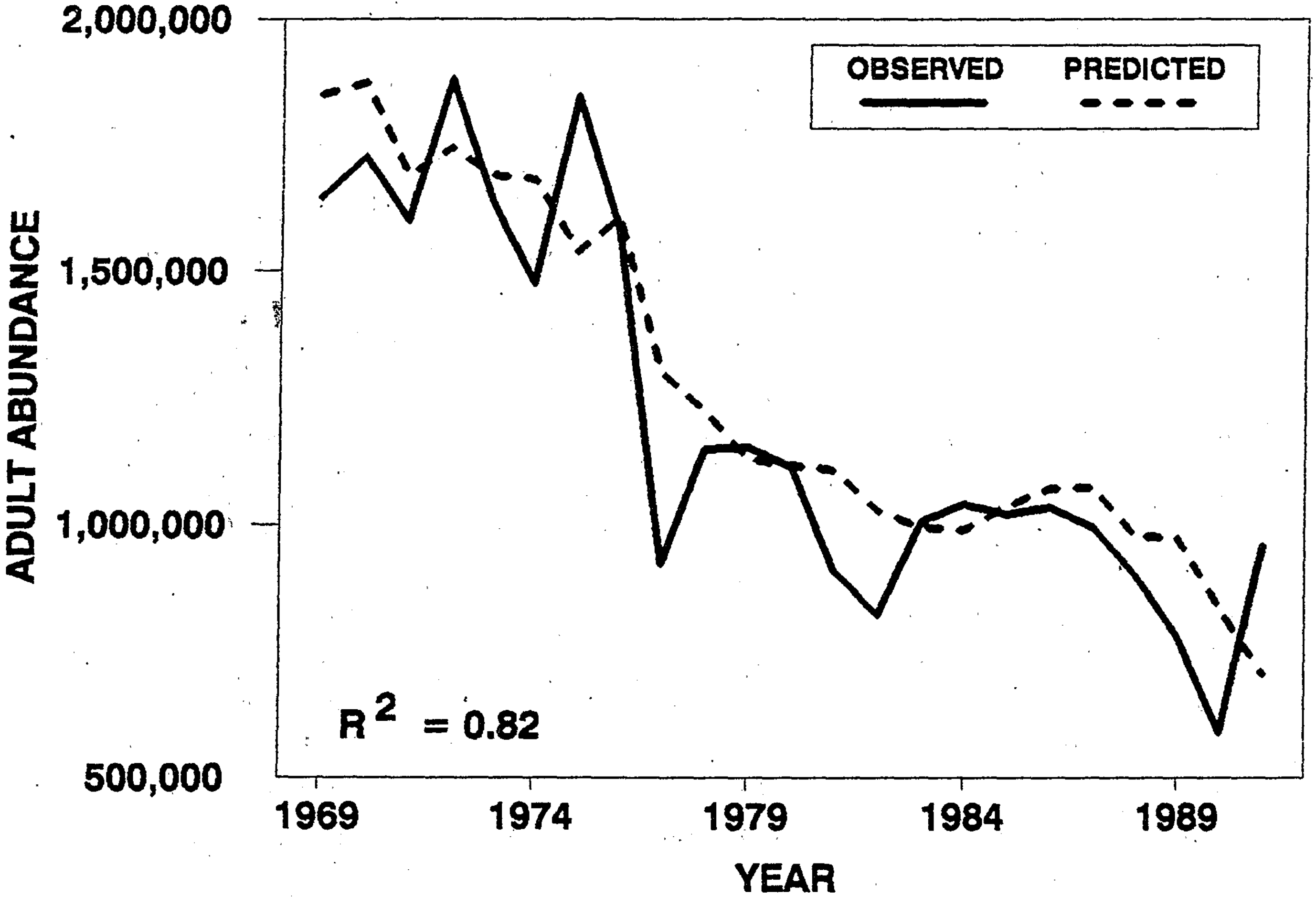



Figure 14. Observed and predicted adult striped bass abundance where predicted values are based on April-December outflow, April-March exports, and adult stock size.

Table 15 produces the same number of fish each year by balancing initial populations (as measured by the yoy index) with export loss rates after the index is set. Thus, low initial abundance requires a reduction in loss rate to produce the same numbers of adults as high initial abundance produces with a high loss rate.

The sensitivity of the output variable in the model, sustained adults, to proportional changes in each of the input variables (initial adults, April-July outflow, August-December outflow, April-July exports, and August-December exports) was evaluated by increasing or decreasing each of the input variables by various percentages and determining the percentage change in sustained adults. Results of this sensitivity analysis suggest that changes in April-July outflow have substantially more effect in dry than in wet year types and that changes in fall and winter water export have greater impact on adult striped bass abundance in wet years (Table 16). Changes in fall-winter export have proportionally more impact than changes in spring and early summer export. This differential in effect between spring and fall-winter exports is greatest in dry years with lower initial adult abundance. The effect of changes in initial adult bass abundance is greater than any of the environmental variables when adult abundance is high.

It is important to recognize that the values in Table 16 underestimate the true impact of the proportional changes in flows and exports if they were sustained over enough years so that they continued to affect the population after it responded as shown in the table. The alterations in egg production associated with the population changes would result in continued population increases or decreases until new equilibriums were reached.

## REFERENCES

California Department of Fish and Game (CDFG). 1992. Revised and updated estimates of fish entrainment losses associated with the State Water Project and Federal Central Valley Project facilities in the south Delta. WRINT-DFG-Exhibit 1 submitted to the State Water Resources Control Board, June, 1992.

Interagency Ecological Study Program for the Sacramento-San Joaquin Estuary (IESP). 1987. Factors affecting striped bass abundance in the Sacramento-San Joaquin Estuary. IESP Technical Report 20:144p + Appendices.

Kano, R. M. 1985. 1984 Clifton Court Forebay evaluations of predation losses to juvenile chinook salmon and striped bass. Memorandum to Clifton Court Forebay Files dated 3/18/85. California Department of Fish and Game, 9p.

Kano, R. M. 1986. 1986 evaluations of predation losses to juvenile striped bass. Memorandum to Clifton Court Forebay Files dated 8/28/86. California Department of Fish and Game, 5p.

Stevens, D. E. 1977a. Striped bass (*Morone saxatilis*) monitoring techniques in the Sacramento-San Joaquin Estuary. In: W. VanWinkle, ed. Proceedings of the conference on assessing the effects of power-plant-induced mortality on fish populations. Pergamon Press, New York, New York, p. 91-109.

Stevens, D. E. 1977b. Stripe bass (*Morone saxatilis*) year class strength in relation to river flow in the Sacramento-San Joaquin Estuary, California. Transactions of the American Fisheries Society 106:34-42.

Stevens, D. E., D. W. Kohlhorst, L. W. Miller, and D. W. Kelley. 1985. The decline of striped bass in the Sacramento-San Joaquin Estuary, California. Transactions of the American Fisheries Society 114:12-30.

Turner J. L. and H. K. Chadwick. 1972. Distribution and abundance of young-of-the-year striped bass, *Morone saxatilis*, in relation to river flow in the Sacramento-San Joaquin Estuary. Transactions of the American Fisheries Society 101:442-452.

The most serious concern I have regarding data subjected to analysis involves calculations used to estimate export losses due to entrainment/predation. Kohlhorst et al. appear to assume that entrained YOY bass suffer a constant 82% predation loss in the SWP's Clifton Court Forebay. This assumption seems logically untenable and appears inconsistent with the 1986 Interagency Report. First, a constant predation loss would not be expected if (a) predator abundance varied, but prey abundance was fixed, or if (b) predator abundance was fixed but prey abundance varied. Only through smooth and implausible joint fluctuations in predator and prey abundance could a constant rate be achieved. Second, the 1986 report, at page 91, states that "predator losses are <u>inversely</u> related to [export] pumping rates". My interpretation of this language is that predation rates would be less under conditions of greater export flows, possibly because duration of YOY bass to predators (primarily adult bass ?) would be decreased. At any rate, I really have no idea how these export loss calculations were made and there are central to the draft CFG impact model. The 1986 document only presents summaries of results of some mark-recapture studies of experimental bass groups released at the "radial gate" and at the "trashboom" of the Clifton Court Forebay.

<u>Statistical Models</u>

As I read the draft report by Kohlhorst et al., they are using regression analyses for two general purposes: (1) to establish statistical relations among (a) adult bass abundance, YOY abundance Indexes, and export "loss rates"; and (2) to establish a connection between "loss rates" and Sacramento water management (export and Delta outflow). Based on these analyses, they then attempt to develop (3) a statistical "management model" whereby export and Delta outflow could be manipulated to produce certain levels of adult striped bass abundance. "Loss rate" is defined as the calculated export losses in year $t$ divided by the YOY index in year $t$.

1. <u>Adult bass abundance vs mean YOY index (3-7 years earlier) and mean loss rate (3-7 years earlier)</u>. Although I am uncertain regarding the general effect of relating adult bass abundance in year $t$ to arithmetic means of YOY indices and loss rates in the previous 3-7 years, I cannot agree that such "error-averaging" across years should generally produce "statistically better results than a relationship simply based on recruitment at age 3 and YOY and losses 3 years earlier" (quotes from p. 11 of Kohlhorst et al.). I also find that arithmetic means are inappropriate because each YOY index should be "discounted" by the survival from year $t$ to year $t+i$, where $i = 3,4,5,6,7$. These survivals from YOY stage to age $i$ would, of course, be progressively smaller, thus suggesting some weighting (as in their refinement 2) at p. 10).

$$\text{Export loss}_t = \alpha \cdot \text{YOY index} \cdot F(\text{export flow, Delta outflow}),$$

where $\alpha$ is a scalar accounting for the unknown relation between true YOY abundance and the YOY Index, and $F(\cdot)$ is an unknown function. Dividing through by the YOY index and taking natural logs gives:

$$ln\ (\text{Export loss/YOY Index}) = ln\ \text{Loss Rate} = ln\alpha + lnF(\cdot)$$

For $F(\cdot) = e^{\beta \text{Export}}$, this would give:

(A) $ln$ Loss Rate = $ln\alpha$ + $\beta$Export

as at middle page 7. If instead $F(\cdot) = e^{\beta \text{Export} + \gamma \text{Delta Outflow}}$, one gets:

(B) $ln$ Loss Rate = $ln\alpha$ + $\beta$Export +$\gamma$Delta Outflow

as at top page 8. Although the authors suggest that forcing model (A) through the origin would prevent non-zero loss rate when Exports are zero (see refinement 1), it is not immediately clear to me that this would be an improvement and it would result in substantial ambiguity regarding interpretation of goodness of fit.

My more substantial concerns with these latter analyses concerns the contention that losses throughout the August-March period must be considered. Although this is probably true at a certain level, it also appears that losses during January-March have nearly always been small when compared to annual losses (with the exception of the 1977 drought year). The authors fail to give adequate details regarding how they selected the months for Export and Outflow that were used in the fitted regression model at the top of page 8. I doubt that a strong case for their choices could be made on the basis of regression $R^2$ or some other "objective" statistical criterion, but I believe that such an objective criterion would be desirable.

4. <u>Use of Statistical Models for Evaluating Outflow and Export Standards</u>. I suspect that the authors used the equation at the top of page 8 to predict loss rate from export and Delta outflow; a model incorporating export and Delta flows, revised to incorporate adult stock, to predict YOY index; and then the equation at the bottom of page 3 to predict resulting adult bass abundance from the predicted YOY index and predicted losses. If so, this procedure would require an initial adult abundance level, as suggested at Table 6. However, the authors do not explicitly state that this is what they did, and they should be forced to do so. If this is indeed what they have done, I am not certainly that it is correct in any event. "Predicted" values of YOY Index and Export Loss Rate are *not* the same as calculated values for a particular year that were used to construct the basic equation at

APPENDIX B

Review Scope of Proposed Work:

*A MEANS OF EVALUATING IMPACTS OF ALTERNATIVE OUTFLOW AND EXPORT CRITERIA ON STRIPED BASS IN THE SACRAMENTO-SAN JOAQUIN ESTUARY*

Report to California Department of Water Resources

Sacramento, CA

Prepared by


Joseph G. Loesch, Ph. D.

HC 1, Box 26

Gloucester Point, VA 23062


December 2, 1991

occurrence of the maximal CPUE also makes it economically attractive. The problems of annual dissimilarities in the growth rate, gear avoidance, emigration, and saltwater encroachment can be very sizable. Use of a maximal CPUE may or may not be an applicable index for YOY striped bass abundance in the Sacramento-San Joaquin Estuary; regardless, the data set should be re-examined. There is no statistically valid reason for including the YOY index in the predictive equation simply because it makes "biological sense".

*Specific comments*

p. 1, par. 2 — The largest declines in adult and juvenile abundance appear to occur almost simultaneously during the 1975-77 period, rather than after the lag that would be expected if the primary cause for the adult decline was decreased juvenile production.

p. 3, par. 1 — To give equal weight to five year classes seems unrealistic, it implies that no adult mortality occurred during these ages. Were other combinations tried, and if so what were the results?

Statistical significance and acceptance levels need be presented in a forthright manner, both in Table 6 and for all subsequent statistical presentations. Including p values would be highly desirable.

Including the non-significant YOY component in the equation is a very questionable procedure, since at all previously observed levels of juvenile abundance the YOY term will make a relatively small contribution to the overall equation and large adult population estimates are possible even if the YOY term is zero. The equation essentially predicts a default population of 1.5 million individuals which can be augmented by up to a few hundred thousand at high levels of juvenile production and which will be linearly depleted by export loss rates, with population extinction inevitable if losses reach about the 1 million mark, which they have in recent years. There definitely seems to be a multi-colinearity problem with the two input variables which could be masking the true effect of juvenile production on ultimate population levels.

The poor fit at the upper end of Figure 8 may be the result of forcing a linear fit to what may be curvilinear relationships. Certainly the relationship in Figure 6 would be expected to pass through the origin and approach an ultimate asymptote, and Figure 7 also suggests a curvilinear relationship.

p. 4, par. 1 — Why are there no observed values for 1966 and 1983 plotted in Figure 9, while they are given in Figure 2? The 1983 value seems to have been ignored, although not obviously omitted, in Figures 11 and 12 as well.

APPENDIX C

UNIVERSITY OF CALIFORNIA, DAVIS

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ

DIVISION OF STATISTICS
STATISTICAL LABORATORY

DAVIS, CALIFORNIA 95616

March 29, 1992

Mr. Jim Sutton
State Water Resources Control Board
Division of Water Rights
P.O. Box 2000
Sacramento, CA 95812-2000

Jim,

I've looked over all of the criticisms that were leveled as the Striped Bass Model that was developed by the Department of Fish and Game. Rather than commenting on the criticisms individually, I found that they subdivide rather nicely into a number of categories, so I'll respond to them categorically instead.

I know that you were hoping to come up with a definitive answer as to whether the striped bass model was *right* or *wrong*. What I *can* say is that the model isn't inherently fallacious, but that there are limitations in the sorts of conclusions that can be drawn from it, some of which are common to all statistical models, and others of which apply particularly to this model. When I make a comment like this, you should bear in mind that I'm a statistician rather than an ecologist by training, and thus I have limited ability to assess how reasonable the assumptions may be on which this model is based.

Quite a few of the criticisms raised in the documents I was provided dealt with technical details of some of the inputs to the model. Since I'm no expert on fisheries or ecology, I can't respond to them. Of the *essentially statistical* comments, I've divided them into four general categories. I'm going to paraphrase each, give a few examples of the type of criticism, and then give my comments on those particular comments:

- *You need to assess the model's accuracy and/or sensitivity to certain inputs.* Chief among these criticisms is the question about the model's sensitivity to the estimated 82% mortality within the Clifton Court Forebay.

  It's certainly true that the value of a statistical model lies both in its ability to provide reasonably accurate predictions of future outcomes, as well as its identification of significant (i.e., influential) factors. Because of this, the statistical significance of a model is only part of the picture it portrays and both its quantitative and qualitative findings will be of interest. In this model, the main qualitative finding is the significance of export in forecasting the loss rate. The quantitative findings lie in the predicted response of the striped bass population to various types of rainfall years and water export strategies. The simplest of these questions to address is which of the factors are significant. Beyond that, the model could perform at any most of the additional predictors that have been suggested also vary with time, and so it's rather difficult to separate between an effect due to water exports and due to other variables that vary similarly, such as the state's population, the number of registered cars, or the national debt, just to name a few that *haven't* been suggested for inclusion in this model. The significance of a given term in a regression model can be viewed only within the context of the other variables that are included in the model. Thus, you can't say definitively that a given variable or set of variables is important, regardless of what else might be put into the model, but rather just that a given variable is important in the context of the particular model in question.

Because there are countless variables that *might* be included in a model like this, I'm more than a little hesitant to play this type of game unless it's been demonstrated that a model including the new variables outperforms the old model, or unless there are biological reasons for choosing the new set of variables instead of the old set. Even if you change around the predictors that are included in the model, this won't necessarily alter the conclusions that come from the model. I'll have more to say about this later on when I discuss the problems associated with trying to impute a causal interpretation to this type of model.

The second aspect of this problem that makes prediction difficult is that the conditions in which we currently find ourselves are in no way similar to the bulk of the data based on which the model was fit. Thinking wishfully, we're coming out of an extended drought, and for whatever reason, the state's fish population has been depleted down to unprecedentedly low levels. It's well known that regression models perform best in the body rather than the extremes of the data, and yet we find ourselves having to make forecasts starting from those extreme conditions. From a statistical standpoint, there's limited (Fisher) information available on which to base those forecasts, and consequently you have to set your sights somewhat lower about this *or any* model's accuracy. Legitimate conclusions can be drawn from the model, such as that the fish population in the next few years will be extremely low, and that it will be lower still if water exports are maintained at elevated levels, but it's unrealistic to expect that you'll get accurate forecasts about just *how low* the population numbers will be. The information on which to base such forecasts simply doesn't exist.

- *The model gives silly (negative) predictions.*

Another manifestation of the problem of drawing inferences for extreme values of the predictor variables is that the slightest misspecification in the model can result in both inaccurate and biased forecasts. This can easily result in negative predictions, but rather than throwing away the entire model because it can predict a negative fish population, you should pay careful attention to the model because it's forecasting *really* low fish numbers. I have to admit that if I had been formulating this type of model, I probably would have chosen the logarithm of the fish index as a dependent variable because many ecological processes are well fitted by lognormal probability models, and because I view the thinning of the fish population as being basically a multiplicative process with random proportions of the population being eliminated at various stages along the way to adulthood. This would have eliminated the problem with negative population estimates, and I think it would also have been more in





C-043034

in a wet year you can do more good for the population than you can *possibly* make up for in a dry year. Moreover, in a wet year, water conservation measures (limits on exports) will be less painful to carry out than in a dry year. That being the case, it makes sense to me to try to beef up the fish population during wet years by restricting the level of water exports, so that the population will be able to withstand the (hopefully only) occasional dry years. I should point out that this last comment is predicated on the fact that the fish population has been restored to reasonable levels. Obviously, the current fish numbers indicate that the population is seriously threatened and as things stand, we can't afford to wait for a wet year to restore the population numbers.

I hope that my comments are useful to you in interpreting the striped bass model. If my comments seem negative in tone, that wasn't my intention. However, I thought it was important to point out what a statistical model can reasonably be expected to accomplish and what it can't.

Sincerely,

Neil H. Willits

Neil H. Willits
Senior Statistician,
Division of Statistics





C-043035